基于 $L_{1/2}$ 正则化和局部纹理约束的人脸超分辨率图像重建 *

王宝成[1, 2]，李 波[1, 2]

(1. 武汉科技大学 计算机科学与技术学院，武汉 430065; 2. 智能信息处理与实时工业系统湖北省重点实验室，武汉 430065)

摘 要：为了更好地利用低分辨率人脸图像，提出了一种基于 $L_{1/2}$ 正则化和局部纹理约束的人脸超分辨率重建。在人脸重建的过程中，使用非负矩阵分解将人脸图像放大到合适的中等分辨率，使用局部纹理约束加强纹理特征提取；然后使用局部稀疏先验进行人脸图像重建，再次加入重建约束和局部纹理约束。为了使获取的人脸图像稀疏系数更加稀疏，使用 $L_{1/2}$ 正则化求解稀疏表示系数。实验结果表明，重建的人脸图像保持了原图像的结构，可以实现很好的重建结果，并且具有更好的鲁棒性。

关键词：稀疏表示；人脸图像；图像重建；$L_{1/2}$ 正则化；局部纹理约束

中图分类号：TP391.41 doi: 10.19734/j.issn.1001-3695.2018.06.0578

Face image super-resolution via $L_{1/2}$ regularization and local texture constraints

Wang Baocheng[1, 2], Li Bo[1, 2]

(1. *College of Computer Science & Technology, Wuhan University of Sciences & Technology, Wuhan* 430065, *China*; 2. *Hubei Province Key Laboratory of Intelligent Information Processing & Real-time Industrial System, Wuhan* 430065, *China*)

Abstract: In order to make better use of low-resolution face images, a new face super-resolution reconstruction based on $L_{1/2}$ regularization and local texture constraints is proposed. In the process of face reconstruction, this paper uses Nonnegative Matrix Factorization to amplify the face image to an appropriate medium resolution. In this process, local texture constraints are used to enhance texture feature extraction. Then, face image reconstruction uses local sparse prior, and adds reconstruction constraints and local texture constraints again. In order to make the obtained face image sparse coefficient more sparse, $L_{1/2}$ regularization is used to solve the problem. The experimental results show that the reconstructed face image retains the original image structure, and it can achieve good reconstruction results, and has better robustness.

Key words: super representation; face image; image reconstruction; l1/2 regularization; local texture constraint

0 引言

目前，超分辨率重建(super-resolution，SR)技术被广泛的运用在遥感、医学成像和图像处理等方面。SR 的目标就是从一个或者多个低分辨率(low resolution, LR)观测图像重建出高分辨率(high resolution，HR)图像。

传统的 SR 方法总是需要同一场景下的多个 LR 图像，但是在大多数的情况下，无法获得多个 LR 图像。因此，单帧图像超分辨率重建(single image super resolution, SISR)获得越来越多的关注。一般来说，由于 LR 图像含有的信息不足，未知的下采样算子和模糊算子，造成了 SISR 是一个“ill-posed”问题，即造成了重建图像的不确定性[1]。在一些传统的重建方法中，如双线性插值[2]和双三次插值[3]，这些简单的插值方法通常会产生具有振铃和锯齿状的过度平滑的图像[4]。近年来，Chang 等人[5]采用流行学习中的局部线性嵌入(LLE)[6]的思想，假设高分辨率和低分辨率补丁空间中的两个流行之间的相似性生成高分辨率图像块。然而，使用固定数量的 K 近邻进行重建，往往会导致模糊效应。Yang 等人[7]基于压缩感知理论，提出一种通过超分辨率重建技术对图片进行重建的方法，该方法在 LR 和 HR 块以及其对应的字典具有相同的稀疏表示的情况下自适应的选择最相关的重建邻域，重建的 HR 图像就是由稀疏表示系数和学习的 HR 字典生成。后来，Yang 等人[8]通过训练紧凑的字典对，来对这种方法进行改进，减少训练字典的时间。在进行稀疏重建的时候，Yang 使用的是 L_1 正则化来获得稀疏系数。在此基础上，Cao 等人[9]使用 $L_p(0<p<1)$正则化替换 L_1 正则化，获取更加稀疏的稀疏系数。Dong 等人[10]提出了使用块坐标下降优化技术的非负字典学习算法。为了提高非负稀疏编码的准确性，利用学习稀疏编码的空间相关性，提出一种基于聚类的结构化稀疏编码方法。之后，Huang 等人[11]利用神经网络，提出一种完全卷积神经网络，称为双向递归卷积神经网络，用于多帧图像超分辨率重建。

本文在稀疏表示的图像超分辨率重建的基础上，对单帧人脸图像进行重建。使用局部二值模式(local binary patterns , LBP)[12]对人脸进行两次纹理特征提取，实现人脸重建过程中的局部纹理约束，以重建出更加清晰的人脸图像。考虑到 $L_{1/2}$ 正则化对于非凸优化问题具有更好的解决方式，使用 $L_{1/2}$ 正则化替换 L_1 正则化，用于获得更加稀疏的稀疏表示系数。利用迭代加权最小二乘法(iterative reweighted least squares, IRLS)进行转换，使 $L_{1/2}$ 正则化适用于非凸优化问题。实验表

收稿日期：2018-06-25；**修回日期：**2018-08-08 **基金项目：**国家自然科学基金资助项目（61572381）

作者简介：王宝成（1992-），男，湖北襄阳人，硕士研究生，主要研究方向为模式识别、人工智能、机器学习（1623956251@qq.com）；李波（1975-），男，教授，博士，主要研究方向为机器学习、模式识别、图像处理.

明，相对于使用 L_1 正则化，使用 $L_{1/2}$ 正则化的超分辨率重建能够获得更稀疏的稀疏系数，而且可以实现比其他重建方法实现更好的重建效果。

1　L1/2 正则化的图像超分辨率重建

在基于稀疏表示的单帧图像超分辨率重建中，输入信号 $x \in R^n$ 可以表示为高分辨率字典 D_h 的稀疏线性组合：$x = D_h\alpha$，α 为 x 的稀疏系数，在实验中可以观察到的人脸图像为 $y = Lx = LD_h\alpha$，其中 L 为投影矩阵。对于未知系数 α，$x = D_h\alpha$ 是不确定的，$y = LD_h\alpha$ 就更加不确定，这就造成了超分辨率重建的“ill-posed”问题。为了约束这个“ill-posed”问题，添加两个条件：

a)重建约束。重建的高分辨率图像 X 可以重新得到输入的低分辨率图像 Y：

$$Y = HBX \tag{1}$$

H 表示下采样算子, B 表示模糊算子。

b)稀疏先验。高分辨率图像 X 的图像块 x 可以表示为高分辨率图像块字典 D_h 的稀疏线性组合：

$$x{=}D_h\alpha \quad \text{for some } \alpha \in R^K \text{ with } \|\alpha\|_0 \ll K \tag{2}$$

稀疏表示 α 通过低分辨率图片 Y 的图像块 y 来进行求解:

$$\min\|\alpha\|_1 \quad s.t.\|FD_l\alpha - Fy\|_2^2 \le \varepsilon \tag{3}$$

其中：ε 为一个很小的误差。考虑到相邻图像块之间的兼容性，最终形成优化问题:

$$\min\|\alpha\|_1 \quad s.t.\begin{matrix}\|FD_l\alpha - Fy\|_2^2 \le \varepsilon_1 \\ \|PD_h\alpha - s\|_2^2 \le \varepsilon_2\end{matrix} \tag{4}$$

其中：矩阵 P 提取当前目标图像块与之前重建的高分辨率图像之间的重叠区域，s 包含重叠的高分辨率图像的像素值。由拉格朗日乘数合并式(4)为

$$\min_{\alpha}\|\tilde{F}\alpha - \tilde{y}\|_2^2 + \lambda\|\alpha\|_1 \tag{5}$$

在文献[13]中证明，$L_{1/2}$ 正则化会产生比 L_1 正则化更加稀疏的解，而且求解过程要比 L_0 正则化更加简单。因此，使用 $L_{1/2}$ 正则化代替 L_1 正则化。则式(5)的最终形式为

$$\min_{\alpha}\left\|\tilde{F}\alpha - \tilde{y}\right\|_2^2 + \lambda\|\alpha\|_{1/2} \tag{6}$$

其中：$\tilde{F} = \begin{bmatrix} FD_l \\ PD_h \end{bmatrix}$，$\tilde{y} = \begin{bmatrix} Fy \\ s \end{bmatrix}$。考虑到细节信息，文献[9]利用 IRLS 法转换式(6)的最优化问题。由于 IRLS 方法通常基于 L_p，p≥1，对于本文中 p=1/2 的情况需要进行转换，使得 IRLS 方法有效的作用于非凸优化问题。则(6)最终转换为新的目标函数:

$$\min_{\alpha}\left\|\tilde{F}\alpha - \tilde{y}\right\|_2^2 + \lambda\sum_{j=1}^{n} w_i\alpha_{ij}^2 \tag{7}$$

其中：$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$，$w_i$ 为前一次 α_{ij}^{k-1} 迭代所得的权值，根据 IRLS 方法，推出如下具体迭代形式:

$$w_j^k = \left[\left(\alpha_{ij}^{k-1}\right)^2 + \varepsilon_3\right]^{-3/4} \tag{8}$$

其中：ε_3 为一个很小的误差。求出式(7)的最优解 α^*，那么高分辨率图像块可以被重建为 $x = D_h\alpha$。之后高分辨率图像块整合为重建的高分辨率图像 X_0。通过将图像 X_0 投影到重建约束条件 $Y = HBX$ 的解空间中来减小误差:

$$X^* = \arg\min_{X}\|HBX - Y\|_2^2 + \varphi\|X - X_0\|_2^2 \tag{9}$$

求解出式(9)的最优解，即为重建的高分辨率图像。

2　局部纹理约束的人脸超分辨率重建

对于人脸图像超分辨率重建，基本的思想是首先把人脸图像放大到合适的中等分辨率图像，然后使用局部稀疏先验模型来重建高分辨率图像块，主要分为两步：a)全局模型,使用重建约束来恢复中等高分辨率人脸图像，但是在人脸子空间中搜索解决方案；b)局部模型,利用稀疏先验来重建高分辨率图像。

2.1　非负矩阵分解

非负矩阵分解(nonnegative matrix factorization, NMF)[14]: 在人脸图像超分辨率重建中，最常用的子空间方法是主成分分析(PCA)，但是由于 PCA 的整体性，PCA 往往难以解释。因此在本文中使用 NMF。NMF 用来寻求给定的信号的表示作为局部特征的组合，并形成以下优化问题：

$$\arg\min_{U,V}\|X - UV\|_2^2 \quad s.t. U \ge 0, V \ge 0 \tag{10}$$

其中: $x \in R^{n \times m}$ 是数据矩阵，$U \in R^{n \times r}$ 是基础矩阵，$V \in R^{r \times m}$ 是系数矩阵。

人脸的超分辨率重建问题可以使用重建约束表示为:

$$c^* = \arg\min_{c}\|HBUc - Y\|_2^2 + \eta\|\delta Uc\|_2 \quad s.t. c \ge 0 \tag{11}$$

其中: $c \in R^{r \times 1}$ 是系数向量，η 为惩罚系数，δ 表示执行高通滤波的矩阵。考虑到局部的纹理信息，以实现更好的重建效果：

$$\begin{gathered} c^* = \arg\min_{c}\|HBUc - Y\|_2^2 + \eta\|\delta Uc\|_2 \\ s.t.\ c \ge 0 \ and\ \|\phi Uc - \phi Y\|_2^2 \le \varepsilon_4 \end{gathered} \tag{12}$$

由拉格朗日乘数可得

$$c^* = \arg\min_{c}\|HBUc - Y\|_2^2 + \eta\|\delta Uc\|_2 + \lambda\|\phi Uc - \phi Y\|_2^2 \tag{13}$$

其中: λ 是惩罚系数，ϕ 是特征提取算子，2.2 节中说明。求解出其最优化解 c^*，那么得到中等分辨率图像为 $\tilde{x}{=}Uc^*$。对于得到的中等分辨率图像 $\tilde{x}$ 的图像块 y，求解出优化问题(7), 得到稀疏系数 α，则高分辨率图像 $X_0{=}D_h\alpha$，得到的高分辨率图像 X_0 要满足重建约束，即要满足式(9): $X^* = \arg\min_{X}\|HBX - Y\|_2^2 + \varphi\|X - X_0\|_2^2$。在对式(9)的求解过程中，再次加入纹理约束:

$$X^* = \arg\min_{X}\|HBX - \tilde{x}\|_2^2 + \mu\|X - X_0\|_2^2 + \gamma\|\phi X - \phi\tilde{x}\|_2^2 \tag{14}$$

其中：$\tilde{x}$ 是中等超分辨率图片，作为新的低分辨率图片，X_0 是重建之后的高分辨率图片，γ 为惩罚系数。求解出式(14)的最优解 X^*,即为最终重建的人脸高分辨率图像。

2.2　局部二值模式

局部二值模式(local binary patterns , LBP)[12]：在进行纹理特征约束的过程中，即局部特征提取的过程中，使用的是局部二值模式。

局部二值模式是机器视觉领域中用于描述图像局部纹理特征的算子，具有旋转不变性和灰度不变性等显著的优点。局部二值模式是一个简单但非常有效的纹理运算符，其最重要的属性是对光照变化造成的灰度变化的鲁棒性。提取 LBP 的过程首先是将原始图像转换为 LBP 图，然后统计 LBP 图的 LBP 直方图，并以这个向量形式的直方图来表示原始的图像。LBP 可以定义为

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p s\left(i_p - i_c\right) \tag{15}$$

其中：(x_c, y_c) 是中心点的像素，亮度为 i_c；而 i_p 是相邻像素的亮度。*S* 是一个函数:

$$s(x)=\begin{cases}1 & if \quad x\geq 0\\ 0 & else\end{cases} \tag{16}$$

然后计算图像块的直方图，即每个数字出现的频率，之后对该直方图进行归一化处理。最后将得到的统计直方图连接成一个特征向量，也就是整幅图的 LBP 纹理特征向量，使用生成的 LBP 纹理特征向量对重建的超分辨率图片进行纹理约束。

3 实验结果

在本文的实验中，实验的数据来自于两个公共人脸数据库：FEI face database 和 CMU+MIT face database 数据库[15]，实验环境为 i3-4160 CPU 3.50 GHz,4.00 GB RAM 的 PC，算法在 MATLAB R2014a 运行。FEI face database 是巴西人脸数据库，是 FEI 人工智能实验室拍摄的一组人脸图像，所有的图像都是彩色的，所有面孔主要由 FEI 的学生和工作人员代表，具有独特的外观、发型和装饰，包含了约 100 个男性和 100 女性测试者共 200 人的数据，每个人有两张人脸图像：一张正常表情图像，一张具有微笑表情的图像。CMU+MIT face database 包含了近 130 幅图像，近 500 个正面人脸图像。

示例样本如图 1 所示，测试结果的时候，使用一个人的两个表情，以验证提出方法的有效性；同时在实验的时候，对人脸图像加入噪声，以验证提出方法对噪声的鲁棒性。在重建过程中，选择低分辨率图像块的大小为 3×3，中等分辨率图像块的大小为 6×6，最终生成的高分辨率图像块的大小为 9×9。最后把重建得到的高分辨率图像块进行整合，形成完整的高分辨率人脸图像。

3.1 不含噪人脸图像重建

对不含噪声的图片进行重建，本文设置式(7)中的参数 λ=0.01。图 2 为重建的结果，表 1 为对应的人脸图像的重建参数，本文选择峰值信噪比 (peak signal to noise ratio，PSNR) 和结构相似性 (structural similarity index，SSIM)作为定性分析的结果。

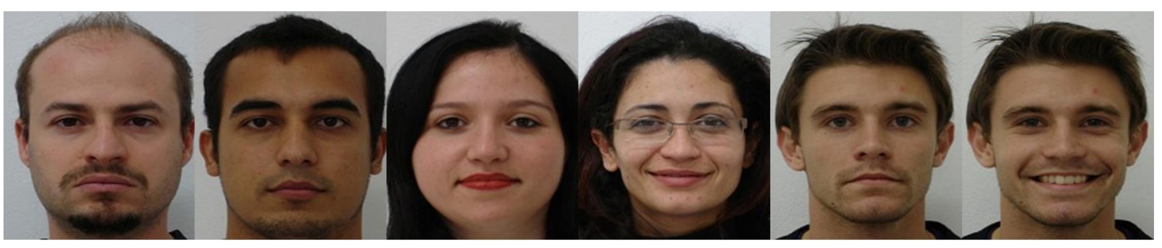


图 1 实验中的高分辨率人脸图像

Fig. 1 High resolution face images in experiments



图 2 图像重建结果

Fig. 2 Image reconstruction results

在图 2 中，图片从左到右依次为分别为低分辨率人脸图像，双三次插值[3]、稀疏表示[8]、领域嵌入(neighbor embedding, NE)[5]、L_1 正则化的局部纹理约束重建和 $L_{1/2}$ 正则化的局部纹理约束重建。实验对两个人的两个表情进行重建，以验证提出方法对人脸微笑表情的鲁棒性。从重建的结果中可以看出双三次插值方法得到的人脸图像趋于平滑和锯齿状伪影，这也说明在表1中双三次插值的PSNR比其他方法较差的原因；稀疏表示重建的图像人脸轮廓较为清晰，但是在较低分辨率人脸图像重建的过程中会出现一些小的伪影；NE 方法可以产生良好的边缘，但是会使得重建的人脸图像趋于平滑,一些特征点,比如人脸三和人脸四脸上的斑纹，没有重建出来，这从表 1 中的 PSNR 参数中也说明 NE 方法相对于稀疏表示来说，在重建结果上会有些失真；相对于 L_1 正则化，$L_{1/2}$ 正则化重建的图片在视觉效果上面无法很好的区分，但是从表 1

中的参数 PSNR 和 SSIM 来看，$L_{1/2}$ 正则化都优于 L_1 正则化重建的结果，这也说明使用 $L_{1/2}$ 正则化求解非凸优化问题可以产生比 L_1 正则化更加稀疏的解。

表 1　不同方法的实验结果对应的 PSNR 和 SSIM

Table 1　PSNR and SSM corresponding to different methods

图片	参数	NE	双三次插值	稀疏表示	L_1	$L_{1/2}$
人脸一	PSNR	35.3672	33.3834	35.0539	37.8949	**38.1789**
	SSIM	0.905311	0.9197	0.8936	0.9307	**0.9353**
人脸二	PSNR	35.2827	33.3738	34.9596	37.9529	**38.1583**
	SSIM	0.8962	0.9184	0.8973	0.9310	**0.9358**
人脸三	PSNR	34.0921	32.9551	34.1616	37.1721	**37.3405**
	SSIM	0.8676	0.8876	0.8847	0.9057	**0.9106**
人脸四	PSNR	33.2124	32.2137	33.9972	36.9143	**36.9558**
	SSIM	0.8598	0.8764	0.8791	0.8975	**0.9007**

3.2　含噪人脸图像重建

在实际生活中，获得到人脸图像，都会有些噪声的影响。为了验证提出方法对含噪的人脸图像的鲁棒性，本文在四张人脸图像中加入均值为 0，方差 σ=2,4,6 的高斯噪声。文献[9]中的研究证明，参数 λ 与噪声方差 σ 正相关，噪声越大，λ 越大。因此对于上面的噪声 σ 设定，本文分别设置 λ=0.2,0.4,0.6，以处理相对应的噪声。表 2 中显示不同方法对不同的噪声重建结果的 PSNR。从表中可以看出，随着噪声的增大，所有方法产生的 PSNR 都在减小,但是 NE 方法下降的最快，双三次插值其次。而稀疏表示和 L_1 正则化和提出的方法下降趋于稳定，而且 L_1 正则化方法随着噪声的增大，其重建结果的 PSNR 优于其他方法的结果。而本文提出的方法通过对噪声的处理，对噪声有一定的鲁棒性，其重建的结果产生的 PSNR 也一直优于其他方法，PSNR 的下降趋势也相对稳定，这些结果也证明了提出方法对噪声的鲁棒性。

表 2　不同方法对不同噪声凸显重建结果的 PSNR

Table 2　PSNR of different reconstruction methods on different noise saliency

噪声	方法	人脸 1	人脸 2	人脸 3	人脸 4
σ=2	NE	33.7829	33.7596	32.7602	32.1332
	双三次插值	31.4888	31.4645	31.2419	30.7529
	稀疏表示	31.4183	31.4732	31.2954	31.1335
	L_1 正则化	34.3739	34.1449	33.8299	33.8510
	$L_{1/2}$ 正则化	**34.6785**	**34.6689**	**34.3436**	**34.3026**
σ=4	NE	30.4482	30.5189	29.9450	29.7752
	双三次插值	28.4796	28.3725	28.2992	28.0788
	稀疏表示	29.5720	29.5778	29.4271	29.4398
	L_1 正则化	30.9482	31.1471	31.0550	30.9690
	$L_{1/2}$ 正则化	**32.4489**	**32.7009**	**32.3522**	**32.2231**
σ=6	NE	27.7324	27.6277	27.4430	27.1214
	双三次插值	25.9815	25.8159	25.7056	25.6607
	稀疏表示	28.7784	28.6308	28.7983	28.6999
	L_1 正则化	29.7794	29.6746	29.7503	29.6423
	$L_{1/2}$ 正则化	**31.1517**	**30.8364**	**30.6793**	**30.9619**

4　结束语

本文提出了一种新的基于 $L_{1/2}$ 正则化和局部纹理约束的人脸超分辨率图像重建。在人脸重建的过程中，先使用非负矩阵分解将低分辨率人脸图像放大到中等分辨率人脸图像，在非负矩阵分解的过程中，使用一次局部纹理约束，加强局部细节特征提取；在中等分辨率重建为高分辨率人脸图像的过程中，再次使用局部纹理约束；同时，为了获得更加稀疏的解，使用 $L_{1/2}$ 正则化代替 L_1 正则化，这样不仅保留了稀疏表示重建的优势，而且可以获得更加稀疏的解，同时可以加强人脸局部特征的提取。从实验结果上面来看，重建后的人脸图像无论是从视觉效果上还是参数分析上，都优于其他的实验结果，证明了提出方法的有效性。

参考文献：

[1] Yang Shuyuan, Wang Min, Sun Yaxin, *et al*. Compressive sampling based single-image super-resolution reconstruction by dual-sparsity and non-local similarity regularizer [J]. Pattern Recognition Letters, 2012, 33 (9): 1049-1059.

[2] Andrew C H, Pattern C L. Digital interpolation of discrete image [J]. IEEE Trans on Computers, 1976, C-25(2): 196-202.

[3] Hou H, Andrews H. Cubic splines for image interpolation and digital filtering [J].IEEE Trans on Acoustics, Speech, and Signal Processing, 1978, 26(6): 208-517.

[4] Lu Xiaoqiang, Yuan Yuan, Yan Pingkun. Image super-resolution via double sparsity resularized manifold learning [J]. IEEE Trans on Circuits and Systems for Video Technology, 2013, 23(12): 2022-2033.

[5] Chang Hong, Yeung Dit-Yan, Xiong Yimin. Super-resolution through neighbor embedding [C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2004: I-275-I-282.

[6] Wen Ying, Zhou Zhenyu, Wang Xunheng, *et al*. An improved locally linear embedding for sparse data sets [C]//Proc of IEEE International Conference on Image Processing.2010: 1585-1588.

[7] Yang Jianchao, Wright J, Huang T. Image super-resolution as sparse representation of raw image patches [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition.2008: 1-8.

[8] Yang Jianchao, Wright J, Huang T. Image super-resolution via sparse representation [J]. IEEE Trans on Image Processing, 2010, 19(11): 2861-2873.

[9] Cao Feilong, Cai Miaomiao, Tan Yuanpeng. Image super-resolution via adaptive lp (0<p<1) regularization and sparse representation [J]. IEEE Trans on Neural Networks and Learning Systems, 2016, 27(7): 1550-1561.

[10] Dong Weisheng, Fu Fazuo, Shi Guangming, *et al*. Hyperspectral image super-resolution via non-negative structured sparse representation [J]. IEEE Trans on Image Processing, 2016, 25(5): 2337-2352.

[11] Huang Yan, Wang Wei, Wang Liang. Video super-resolution via Bidirectional recurrent convolutional networks [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2018, 40(4): 1015-1028.

[12] Jain A, Rathna G N. Visual speech recognition for isolated digits using discrete cosine transform and local binary pattern features [C]//Proc of IEEE Global Conference on Signal and Information Processing, 2017: 368-372.

[13] Xu Zongben, Zhang Hai, Wang Yao. $L_{1/2}$ regularization [J]. Science China Information Science, 2010, 53(6): 1159-1169.

[14] Wu Zebin, Ye Shun, Wei Jie, *et al*. Parallel optimization of hyperspectral unmixing based on sparsity constrained nonnegative matrix factorization [C]//Proc of IEEE International Geoscience and Remote Sensing Symposium.2013: 1438-1441.

[15] Jiang Junjun, Ma Jiayi, Chen chen, *et al*. Noise robust face image super-resolution through smooth sparse representation [J]. IEEE Trans on Cybernetics, 2017, 47(11): 3991-4002.